Characterizing Extrapolations in Multiple Regression

by

R. Dennis Cook and Sanford Weisberg

Technical Report No. 330

School of Statistics
University of Minnesota
Saint Paul, MN 55108

1 November 1978

This work was supported in part by Grant 1-R01-GM25587-01 from the National Institute of General Medical Science.

Abstract

When a fitted linear regression model is only approximate, valid predictions using the model may be obtained only for cases that fall in some range of applicability. This range should depend on the data used to estimate the regression equation. While precise definition of a range of applicability is impossible without knowledge of the true model, a useful region can be obtained as the smallest convex hull containing the observed data. In this paper, the smallest volume ellipsoid containing this set is suggested as an approximation to it. The ellipsoid can be found using optimal design theory. An algorithm for computation of it is presented, along with two examples.

Key words: Linear models; prediction; minimum containing ellipsoid; optimal designs

1. Introduction

An important use of linear regression models is the prediction of future values. If the fitted linear model is known to be exactly correct, then predictions can be expected to be useful and reliable. In most practical situations, however, a linear model is fit as an approximation to a more complex relationship. Here, useful predictions can be obtained only for new cases that fall in some region of applicability that depends on the data used to estimate the prediction equation. It is well known that extrapolation outside a region of applicability can lead to nonsensical predictions.

Precise definition of a region of applicability for predictions is impossible without complete knowledge of a problem and the true model. However, when the model used is chosen to make the fit adequate for the observed data, a reasonable region is the smallest convex set in the space of the independent variables that includes all of the observed data points; that is, the set of all carrier vectors in the data and all linear combinations of them. Because of the difficulty in describing this set, we propose using the smallest volume ellipsoid that circumscribes the convex set as an approximation to it. We refer to this ellipsoid as the minimum containing ellipsoid (MCE). Once the MCE is known, checking whether or not a point is in the ellipsoid (hence checking if a prediction is reliable) requires only the computation of an inner product.

In Section 2 of this paper we derive the MCE using results from optimal design theory. Section 3 contains two examples and in Section 4 several extensions and related problems are discussed.

2. Minimal Containing Ellipsoid (MCE): Theory

Let $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ denote n points in R^p and assume that these points span R^p. These are the values of the carriers for the n cases in the data to be used to estimate the prediction equation. The MCE problem is to find an ellipsoid in R^p which contains the points $\underline{x}_i$, $i = 1, 2, \ldots, n$, and has smallest volume. A solution to this problem may be found by appealing to optimal design theory. We first briefly review the necessary design theory and then apply it to the MCE problem.

2.1 Optimal Design

Consider the linear model

$$y = \underline{z}'\underline{\beta} + e \tag{2.1}$$

where $\underline{z}$ is a p×1 vector of carriers, $\underline{\beta}$ is a p×1 vector of unknown parameters, and the errors are uncorrelated with mean zero and constant variance. The carrier vectors $\underline{z}$ are to be selected from some compact subset, χ, of R^p. We refer to χ as the design space. An experimental design is specified by a probability measure ξ on χ with the understanding that if ξ places mass π at a point $\underline{z}^* \varepsilon \chi$ then this fraction of the total number of observations is to be taken at $\underline{z}^*$. For experiments of size N, exact designs constrain $N\pi$ to be a non-negative integer while approximate designs are not constrained in this way. Here, we need consider only approximate designs.

Many criteria have been proposed for optimizing the selection of a design, ξ. Generally, they all specify a selection which minimizes some functional of the information matrix, $\underline{M}(\xi)$, defined

$$\underline{M}(\xi) = \int_{\chi} \underline{z}\,\underline{z}'\, d\xi .$$

Designs which minimize the functionals

(a) determinant $(\underline{M}^{-1}(\xi)) = |\underline{M}^{-1}(\xi)|$

and

(b) $\max_{\underline{z} \varepsilon \chi} \underline{z}' \underline{M}^{-1}(\xi) \underline{z}$

are called D- and G-optimal designs, respectively. The following theorem (Kiefer and Wolfowitz, 1960) established the equivalence of these criteria for approximate designs.

Theorem (Equivalence Theorem): The following assertions are equivalent:

(1) The design ξ^* maximizes $|\underline{M}(\xi)|$ (minimizes $|\underline{M}^{-1}(\xi)|$).

(2) The design ξ^* minimizes $\max_{\underline{z} \varepsilon \chi} \underline{z}' \underline{M}^{-1}(\xi) \underline{z}$.

(3) $\max_{z \varepsilon \chi} \underline{z}' \underline{M}^{-1}(\xi^*) \underline{z} = p$.

The information matrices of all designs satisfying (1) - (3) are identical and any convex combination of designs satisfying (1) - (3) also satisfies (1) - (3).

The following geometric interpretation of the Equivalence Theorem was given by Silvey (1972) and is relevant to the MCE problem: Assume that model (2.1) does not contain a constant term and consider ellipsoids, centered at the origin, of the form

$$\underline{z}' \underline{M}^{-1}(\xi) \underline{z} \le p \; . \tag{2.2}$$

Since

$$\int_{\chi} \underline{z}' \underline{M}^{-1}(\xi) \underline{z} \, d\xi = p \; , \tag{2.3}$$

usually such ellipsoids will contain part of χ, but not all of it. However, the Equivalence Theorem shows that there is an ellipsoid of the form (2.2) which has maximum volume and contains (circumscribes) χ: Maximizing volume means maximizing $|M(\xi)|$ which is achieved by assertion (1) while circumscribing χ means having (2.2) hold for all $\underline{z} \epsilon \chi$ and this is achieved by assertion (3). In addition, it follows from (2.3) that the only possible support points for a D-optimal design are those where the circumscribing ellipsoid meets χ.

Silvey (1972) also conjectured that the D-optimal design would provide an ellipsoid of minimal volume out of the class of all ellipsoids which are centered at the origin and contain χ. The proof of Silvey's conjecture was furnished by Sibson (1972). Thus, the inverse of the information matrix of a D-optimal design for a model without a constant term may be used to construct an ellipsoid which has minimal volume out of the class of ellipsoids which are centered at the origin and contain χ. By defining χ to be the set of n points in the MCE problem, $\chi = \{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n\}$, we obtain a partial solution to the problem. A complete solution is found by relaxing the condition that the ellipsoids be centered at the origin. In the next section we show that this is accomplished by forcing model (2.1) to have a constant term.

2.2 Minimal Containing Ellipsoids

Assume that model (2.1) has a constant term and write $\underline{z}' = (1, \underline{w}')$, where now $\underline{w} \epsilon R^p$. The design space χ may still be considered a subset of R^p. Partition the information matrix $\underline{M}(\xi)$ for a design ξ on χ as

$$\underline{M}(\xi) = \begin{pmatrix} 1 & \underline{\mu}'(\xi) \\ \underline{\mu}(\xi) & \underline{C}(\xi) \end{pmatrix}$$

where

$$\underline{\mu}(\xi) = \int_{\chi} \underline{w}\, d\xi \tag{2.4}$$

and

$$\underline{C}(\xi) = \int_{\chi} \underline{w}\,\underline{w}'\, d\xi.$$

Let

$$\begin{aligned}\underline{M}_1(\xi) &= \underline{C}(\xi) - \underline{\mu}(\xi)\underline{\mu}(\xi)' \\ &= \int_{\chi} (\underline{w} - \underline{\mu}(\xi))(\underline{w} - \underline{\mu}(\xi))'\, d\xi \ .\end{aligned} \tag{2.5}$$

Then, it is straightforward to verify that

$$|\underline{M}(\xi)| = |\underline{M}_1(\xi)| \tag{2.6}$$

and

$$\underline{z}'\underline{M}^{-1}(\xi)\underline{z} - 1 = (\underline{w} - \underline{\mu}(\xi))'\underline{M}_1^{-1}(\xi)(\underline{w} - \underline{\mu}(\xi)) \ . \tag{2.7}$$

Let ξ^* denote the D-optimal design. We will show that the p dimensional ellipsoid

$$(\underline{w} - \underline{\mu}(\xi^*))'\underline{M}_1^{-1}(\xi^*)(\underline{w} - \underline{\mu}(\xi^*) \leq p \tag{2.8}$$

is an ellipsoid of minimal volume out of the class of all p dimensional ellipsoids which contain χ. It follows immediately from (2.7) and the Equivalence Theorem that this ellipsoid contains χ. Therefore, it remains to show that (2.8) has minimal volume.

For an arbitrary but fixed point $\underline{d} \in R^p$, let $S(\underline{d})$ denote the set of all $p \times p$ positive definite symmetric matrices, $\underline{N}$, such that

$$(\underline{w} - \underline{d})'\underline{N}(\underline{w} - d) \leq p \tag{2.9}$$

for all $\underline{w} \in \chi$. The set $S(\underline{d})$ indexes the set of all ellipsoids with center at $\underline{d}$ which contain χ. Next, choose $\underline{N}_d \in S(\underline{d})$ such that

$$|\underline{N}_d| \geq |\underline{N}|$$

for all $\underline{N} \in S(\underline{d})$. The matrix $\underline{N}_d$ defines an ellipsoid with minimal volume out of the set $S(\underline{d})$.

In (2.9), replacing $\underline{N}$ with $\underline{N}_d$, replacing $(\underline{w} - \underline{d})$ with $(\underline{w} - \underline{\mu}(\xi^*) + \underline{\mu}(\xi^*) - \underline{d})$ and then integrating both sides with respect to the D-optimal design measure, ξ^*, results in the following inequality:

$$\text{Trace}\left[\underline{M}(\xi^*)\underline{N}_d\right] + (\underline{\mu}(\xi^*) - \underline{d})'\underline{N}_d(\underline{\mu}(\xi^*) - \underline{d}) \leq p \ . \tag{2.10}$$

Next, using the familiar relationship between the arithmetic and geometric means it follows from (2.10) that

$$\left[|\underline{M}(\xi^*)|\cdot|\underline{N}_d|\right]^{1/p} + \frac{(\underline{\mu}(\xi^*) - \underline{d})'\underline{N}_d(\underline{\mu}(\xi^*) - \underline{d})}{p} \leq 1 \ . \tag{2.11}$$

The desired result is obtained from this inequality by noting that since $\underline{N}_d$ is positive definite we must have $|\underline{M}^{-1}(\xi^*)| > |\underline{N}_d|$ for $\underline{d} \neq \underline{\mu}(\xi^*)$ and when $\underline{d} = \underline{\mu}(\xi^*)$, $M^{-1}(\xi^*) = \underline{N}_d$ by Sibson's (1972) duality theorem.

In short, the ellipsoid of minimal volume which contains the points $\underline{x}_1, \ldots, \underline{x}_n$ in R^p may be found by defining the design space $\chi = \{\underline{x}_1, \ldots, \underline{x}_n\}$ and constructing the D-optimal design, ξ^*, for the model

$$y = \binom{1}{\underline{x}}'\underline{\beta} + \varepsilon$$

where $\underline{x} \in \chi$. The MCE is then given by (2.8).

Finally, it is of interest to note that, since the only possible support points for the D-optimal design are where the MCE meets χ, by inspecting ξ^* we can detect at least $p+1$ points which lie <u>on</u> the convex hull of χ.

3. Applications of the MCE

3.1 An iterative method of finding the MCE

Many iterative algorithms for computing D-optimal designs are available. We have used an algorithm given by Federov (1972), p. 102. While this algorithm converges to the D-optimal design monotonically, faster algorithms are possible.

Let $\underline{x}_i$ be the i-th $p \times 1$ data point, $i = 1,2,\ldots, n$. The design ξ at the k-th iteration is specified by a vector of non-negative weights $\underline{\pi}^{(k)} = (\pi_1^{(k)},\ldots\pi_n^{(k)})$, $\Sigma\pi_i^{(k)} = 1$. After k steps of the algorithm, in analogy to (2.5) and (2.6),

$$\underline{\mu}(\underline{\pi}^{(k)}) = \Sigma\pi_i^{(k)}\underline{x}_i$$

$$\underline{M}_1(\underline{\pi}^{(k)}) = \Sigma\pi_i^{(k)}\underline{x}_i\underline{x}_i' - \underline{\mu}(\underline{\pi}^{(k)})\underline{\mu}(\underline{\pi}^{(k)})' .$$

Next compute,

$$m_k = \max_i \left(\underline{x}_i - \underline{\mu}(\underline{\pi}^{(k)})\right)'\underline{M}_1^{-1}(\underline{\pi}^{(k)})\left(\underline{x}_i - \underline{\mu}(\underline{\pi}^{(k)})\right) . \tag{3.1}$$

Now, m_k is essentially the maximum variance at any of the x_i, and if m_k is sufficiently close to p, (e.g. if $m_k < p + \varepsilon$ for ε small), then $\underline{\pi}^{(k)}$ specifies the final design (and iteration terminates). If $m_k > p + \varepsilon$, the design is modified by increasing the weight for the point where the maximum in (3.1) is attained, so that, at the next iteration, the variance at this point is exactly p and this point is then on the boundary of the ellipsoid. Specifically, let

$$\underline{\pi}^{(k+1)} = (1 - \alpha)\underline{\pi}^{(k)} + \alpha\underline{u}$$

where $\underline{u}$ is the $n \times 1$ unit vector with a 1 corresponding to an observation where the maximum in (3.1) is attained and 0 elsewhere, and $\alpha = (m_k - p)/p(m_k - 1)$. Note that at each step both a new inner product matrix $\underline{M}_1^{-1}$ and a new center $\underline{\mu}$ are computed.. The new center is closer to the point where the maximum occured at the last iteration. Let $\underline{\pi}_f$ denote the weights for the design after the final iteration.

For the initial $\underline{\pi}^{(0)}$ it is convenient to choose $\pi_i^{(0)} = 1/n$ $i = 1,\ldots, n$. Then $\underline{\mu}(\underline{\pi}^{(0)}) = \underline{\bar{x}} = (\bar{x}_1, \bar{x}_2,\ldots, \bar{x}_p)'$, $\underline{M}_1(\underline{\pi}^{(0)}) = ((\underline{X}'\underline{X}) - \underline{\bar{x}}\,\underline{\bar{x}}')/n$ is just the corrected cross product matrix divided by n, and the point satisfying (3.1) is the point corresponding to the largest diagonal element of the projection matrix $\underline{V} = \underline{X}(\underline{X}'\underline{X})^{-1}\underline{X}'$. The i-th diagonal entry of $\underline{V}$ is $v_{ii} = ((\underline{x}_i - \underline{\bar{x}})'\underline{M}_1^{-1}(\underline{x}_i - \underline{\bar{x}}) + 1)/n$. The v_{ii} have been recently discussed as measures of leverage or potential influence of individual observation on a regression problem (Hoaglin and Welsch, 1978; Cook, 1977).

3.2 Example 1

As an idealized example with $p = 2$ carriers, a bivariate normal pseudo-random sample of $n = 40$ was generated according to

$$\begin{pmatrix} x_{i1} \\ x_{i2} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 10 & 9 \\ 9 & 1 \end{pmatrix}\right),$$

and then rounded to two decimal places. The data actually obtained are graphed in Figure 1. The solid curve is the ellipse corresponding to the design $\underline{\pi}^{(0)}$. It is centered at the point marked "1" on the graph. The dotted line gives the MCE centered at the point "2" on the graph, obtained after 55 iterations using $\varepsilon = 0.10$. The design that gives MCE is given in Table 1. Notice from Table 1 that evidently 4 points have nonzero weight. Thus four points lie on the convex hull.

As might be expected, little is gained by using the MCE over the ellipse resulting from use of $\underline{\pi}^{(0)}$ in this problem because of the ideal behavior of the data. In more realistic problems, the MCE will represent a substantial improvement over the ellipse determined by $\underline{\pi}^{(0)}$.

To apply the MCE to a prediction at $\underline{x}^*$ we need only compute $m^* = (\underline{x}^* - \underline{\mu}(\underline{\pi}_f))'\underline{M}_1^{-1}(\underline{x}^* - \underline{\mu}(\underline{\pi}_f))$. If $m^* > p$ then $\underline{x}^*$ is not contained in the MCE. For this example, $\underline{\mu}(\underline{\pi}_f)$ and $\underline{M}_1^{-1}(\underline{\pi}_f)$ are given in Table 1.

3.3 Example 2

The data for this example were taken from Draper and Smith (1966, p. 116). There are 25 observations and two predictors, X_1 and X_2, where X_1 is average atmospheric temperature (in ^oF) in a month and X_2 is number of operating days in a month. The data, and the weights for the final iterations ($\varepsilon = 0.10$) are given in the first three columns of Table 2.

Figure 2 gives the ellipses corresponding to $\underline{\pi}^{(0)}$ (solid line) and $\underline{\pi}_f$ (dotted line) obtained after 59 iterations. The centers of the ellipses are indicated as in the previous example. In this case, the ellipses corresponding to $\underline{\pi}^{(0)}$ and $\underline{\pi}_f$ differ considerably. Note also that two points account for a sizable portion of the volume of the ellipse defined by $\underline{\pi}_f$ and that there is a considerable "gap" between these points and the bulk of the data. These points would probably have considerable influence on the final prediction equation and could be detected by other means (Cook, 1977; Cook, 1979; Hoaglin and Welsch, 1978). Depending on the application and knowledge of the true model, predictions within such gaps may not be reliable.

The portion of the volume of the MCE due to the two "outlying" points may be found by deleting these points, recomputing the MCE and calculating $1 - \{|\underline{M}_1(\underline{\pi}_f)| \;/\; |\underline{M}_1(\underline{\pi}_f^*)|\}^{\frac{1}{2}}$, where $\underline{\pi}_f$ and $\underline{\pi}_f^*$ denote the final weights for the complete and reduced data sets, respectively. The vector $\underline{\pi}_f^*$ is given in the fourth column of Table 2 and $\underline{M}_1(\underline{\pi}_f)$, $\underline{M}_1(\underline{\pi}_f^*)$ and their determinants are given at the bottom of Table 2. The initial and final ellipses for the reduced data set are given in Figure 3. In this example, about 67% of the volume of the MCE is due to the two outlying points.

4. Additional Remarks and Extensions

In this paper we have suggested a method for characterizing extrapolations in regression problems that does not depend on the method of estimation. Thus, the region is equally applicable to least squares, ridge, or other estimation techniques.

Higher Order Terms

If certain carriers in a model are functionally related, such as a linear and quadratic term in the same variable, we suggest using only one of the functionally related terms when determining the MCE. Inclusion of all functionally related terms may lead to severe and unnecessary restrictions. For example, consider a model which contains only a linear and quadratic term in a single variable. The MCE for the four points (±9, 81), (±10, 100) would not allow predictions in a neighborhood of the origin.

Subset Selection

Even if some carriers have been eliminated from a prediction equation by a selection procedure, we suggest computing the MCE using all first order carriers available. This results in explicit restrictions on the applicability of the resulting predictions: Consider a first order model and partition the carrier vector $\underline{x}$ as $\underline{x}' = (\underline{x}_1', \underline{x}_2')$. Assume that the coefficients corresponding to $\underline{x}_2'$ have been set to zero by a selection procedure. We suggest that prediction at a point $\underline{x}^*_1$ take the form

$$\hat{y} = \underline{x}_1^{*\prime} \hat{\underline{\beta}}_1$$

for all $\underline{x}_2$ such that $(\underline{x}_1^{*\prime}, \underline{x}_2')$ is in the MCE based on the complete data set.

Multicollinearity

In the presence of multicollinearity in the carriers, some authors, such as Gunst and Mason (1977) have suggested that reliable predictions can be obtained only for new cases lying "near" eigenvectors corresponding to large eigenvalues of the corrected cross product matrix. The MCE serves to make their notion somewhat more precise.

The simplest method of deciding if a point x^* can have a reliably predicted response is to require that each of the p coordinates in x^* be in its observed range, essentially circumscribing the convex hull of the data by a hyper-rectangle. This technique can be expected to be poor if the data exhibits collinearity. However, this method can be combined with the MCE to give a smaller region by requiring reliable predictions to have both x^* in the MCE <u>and</u> each coordinate of x^* to fall in the observed range.

We are grateful to Kinley Larntz and Alan Izenman for their comments, and to Christopher Nachtsheim for programming assistance.

References

Cook, R. Dennis (1977), "Detection of influential observations in linear regression," Technometrics, 19, 15-18.

Cook, R. Dennis (1979), "Influential observations in linear regression," Journal of the American Statistical Association, to appear.

Draper, N.R. and H. Smith (1966), Applied Regression Analysis, New York: John Wiley & Sons, Inc.

Federov, V.V. (1972), Theory of Optimal Experiments, New York: Academic Press.

Gunst, R.F. and R.L. Mason (1977), "Advantages of examining multicollinearities in regression analysis," Biometrics, 33, 249-260.

Hoaglin, D.C. and R.E. Welsch (1978), "The hat matrix in regression and ANOVA," The American Statistician, 32, 17-22.

Kiefer, J. and J. Wolfowitz (1960), "The equivalence of two extremum problems," The Canadian Journal of Mathematics, 12, 363-366.

Sibson, R. (1972), "Discussion on the papers by Wynn and Laycock," Journal of the Royal Statistical Society B, 34, 181-183.

Silvey, S.D. (1972), "Discussion on the papers by Wynn and Laycock," Journal of the Royal Statistical Society B, 34, 174-175.

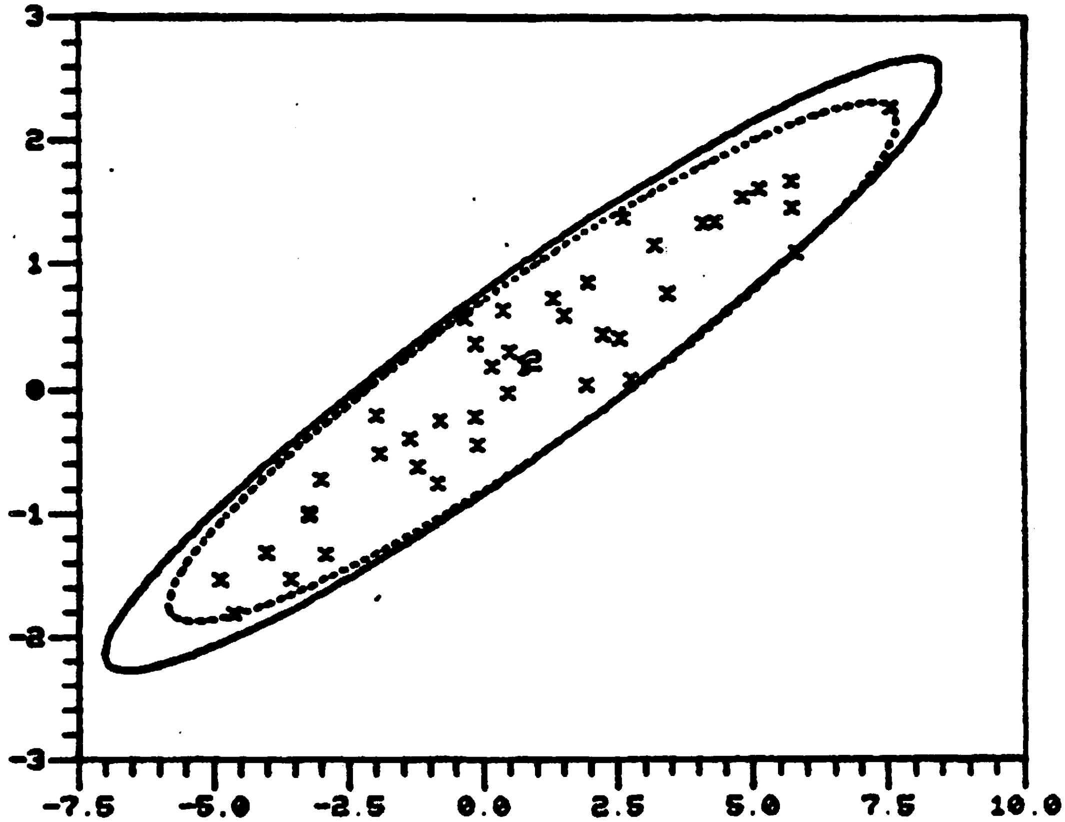

Figure 1. Simulated data, n = 40.

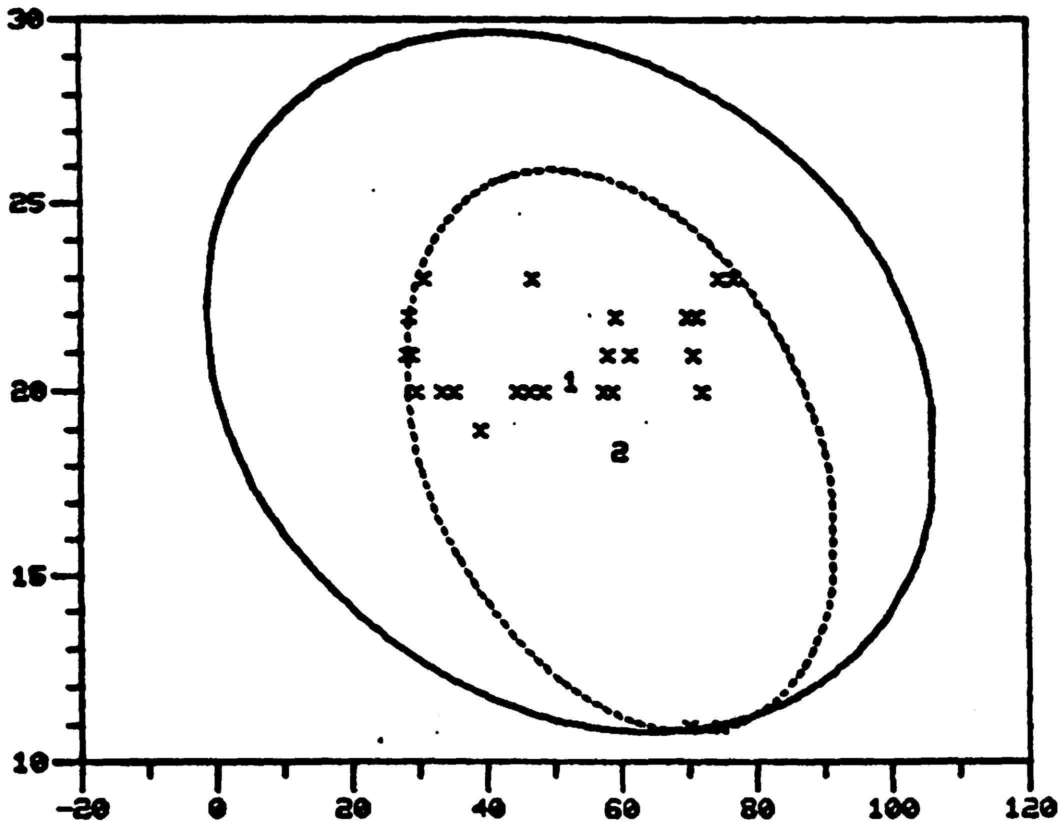

FIGURE 2 — DRAPER AND SMITH EXAMPLE

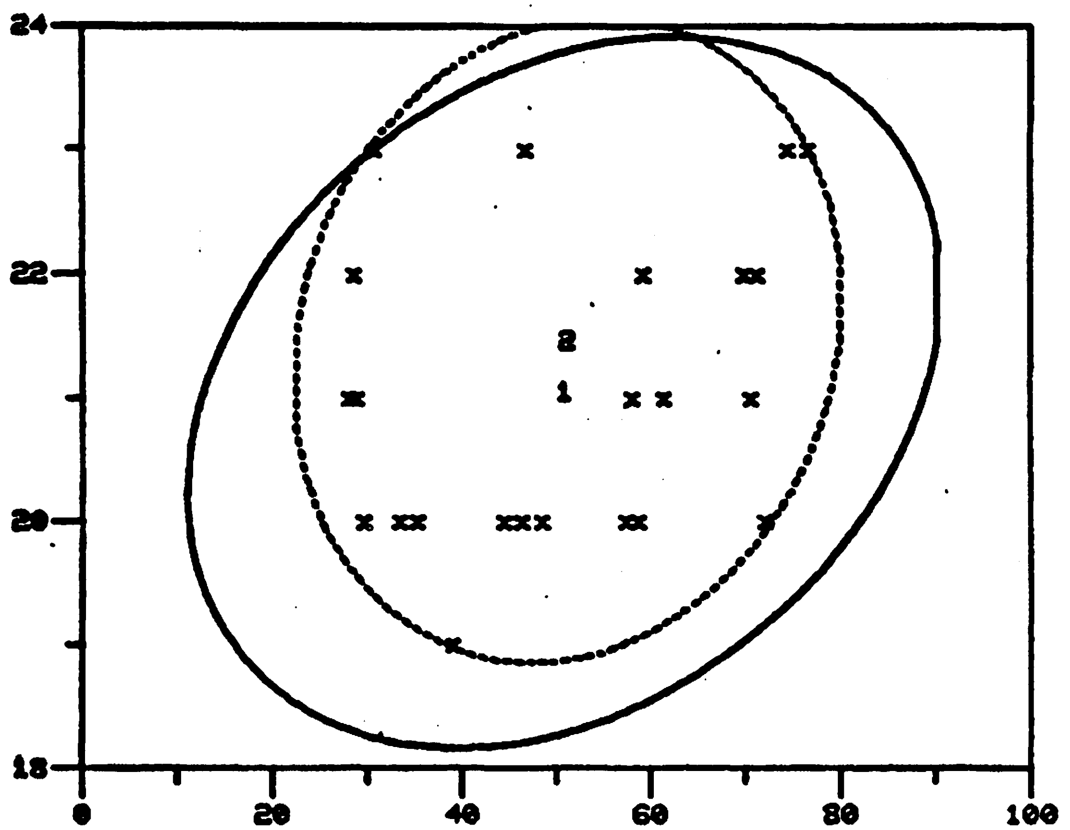

FIGURE 3 — DRAPER AND SMITH EXAMPLE WITH TWO POINTS REMOVED

Table 1. Design and summaries for Example 1

X1	X2	π_f
2.59	1.38	.1120E-02
-1.98	-.18	.1120E-02
1.95	.86	.1120E-02
2.22	.45	.1120E-02
.36	.64	.1120E-02
-4.62	-1.79	.2975
-1.36	-.37	.1120E-02
7.55	2.29	.1439
2.53	.42	.1120E-02
-1.23	-.60	.1120E-02
5.72	1.70	.1120E-02
4.06	1.34	.1120E-02
-3.23	-.98	.1120E-02
.14	.20	.1120E-02
4.30	1.35	.1120E-02
1.50	.60	.1120E-02
5.11	1.63	.1120E-02
.49	.32	.1120E-02
1.92	.05	.1120E-02
-.15	-.19	.1120E-02
-.15	.38	.1120E-02
-3.24	-1.00	.1120E-02
3.42	.77	.1120E-02
-3.02	-.70	.1120E-02
-.86	-.74	.1120E-02
-.34	.58	.2950
1.29	.73	.1120E-02
5.72	1.48	.1120E-02
-3.59	-1.51	.1120E-02
-4.88	-1.52	.1120E-02
-1.91	-.49	.1120E-02
4.81	1.56	.1120E-02
.46	-0	.1120E-02
2.75	.10	.1120E-02
3.19	1.16	.1120E-02
-.10	-.43	.1120E-02
5.81	1.11	.2233
-.79	-.22	.1120E-02
-2.92	-1.31	.1120E-02
-4.03	-1.30	.1120E-02

$$\underline{M}_1^{-1}(\underline{\pi}^{(0)}) = \begin{pmatrix} .909 & -2.759 \\ -2.759 & 9.168 \end{pmatrix} \qquad \underline{M}_1^{-1}(\underline{\pi}_f) = \begin{pmatrix} .351 & -1.057 \\ -1.057 & 3.665 \end{pmatrix}$$

$$\underline{\mu}(\underline{\pi}^{(0)}) = \begin{pmatrix} .7373 \\ .1942 \end{pmatrix} \qquad \underline{\mu}(\underline{\pi}_f) = \begin{pmatrix} .9331 \\ .2224 \end{pmatrix}$$

$\varepsilon = 0.10$, 55 iterations

Table 2. Design and Related Statistics for Data from Draper and Smith

X1	X2	π_f	π_f^*
35.30	20.00	.2073E-02	.2529E-02
29.70	20.00	.2073E-02	.2529E-02
30.80	23.00	.2073E-02	.3004
58.80	20.00	.2073E-02	.2529E-02
61.40	21.00	.2073E-02	.2529E-02
71.30	22.00	.2073E-02	.2529E-02
74.40	11.00	.3254	**
76.70	23.00	.3228	.2529E-02
70.70	21.00	.2073E-02	.1333
57.50	20.00	.2073E-02	.2529E-02
46.40	20.00	.2073E-02	.2529E-02
28.90	21.00	.2073E-02	.2529E-02
28.10	21.00	.3062	.2529E-02
39.10	19.00	.2073E-02	**
46.80	23.00	.2073E-02	.2578
48.50	20.00	.2073E-02	.2529E-02
59.30	22.00	.2073E-02	.2529E-02
70.00	22.00	.2073E-02	.2529E-02
70.00	11.00	.2073E-02	.2529E-02
74.50	23.00	.2073E-02	.2529E-02
72.10	20.00	.2073E-02	.2604
58.10	21.00	.2073E-02	.2529E-02
44.60	20.00	.2073E-02	.2529E-02
33.40	20.00	.2073E-02	.2529E-02
28.60	22.00	.2073E-02	.2529E-02

**Omitted from computations

$$\underline{M}_1^{-1}(\underline{\pi}_f) = \begin{pmatrix} .002298 & .003068 \\ .003068 & .04090 \end{pmatrix} \qquad \underline{M}_1^{-1}(\pi_f^*) = \begin{pmatrix} .002538 & -.003669 \\ -.003669 & .3135 \end{pmatrix}$$

$$\underset{\sim}{\mu}(\underline{\pi}_f) = \begin{pmatrix} 52.60 \\ 20.24 \end{pmatrix} \qquad \underline{\mu}(\pi_f^*) = \begin{pmatrix} 51.23 \\ 21.46 \end{pmatrix}$$

$$|\underline{M}_1(\underline{\pi}_f)| = 11{,}823.7 \qquad |\underline{M}_1(\underline{\pi}_f^*)| = 1{,}278.4$$

In each of the above, $\varepsilon = 0.10$. π_f was reached in 59 iterations, and π_f^* was reached in 54.